CHAPTER IV

POLYHEDRIN PROMOTER BINDING PROTEIN AND ITS COGNATE SEQUENCE RECOGNITION MOTIFs ARE NECESSARY FOR EXPRESSION FROM THE PROMOTER

1. Mutations in the PPBP-cognate sequence motifs within the polh promoter abolish transcription from the polyhedrin promoter in vivo

2. Individual bases within the hexa-motif are sensitive to PPBP binding in vitro

3. Individual bases within the hexanucleotide motif are also important for expression of a luciferase reporter gene from the polyhedrin promoter in vivo

4. PPBP and TATA-binding-protein (TBP) are different with respect to their binding activities

5. PPBP-like activity is present in other eukaryotic systems
In Vivo Role of PPBP-Binding Motifs

The unusual 30-kDa factor, PPBP or the polyhedrin promoter binding protein, binds to a hexa-motif AATAAA and the octa-motif TAAGTATT encompassing the transcription start point (Burma et al., 1994) of the polyhedrin promoter. Interestingly, the 18 bp "minimal" promoter element described by Morris and Miller (1994) essentially consists of the PPBP binding domains. A number of observations suggested that PPBP is important in polyhedrin gene transcription: a) dephosphorylation of PPBP abolishes binding to its cognate sequences within the promoter (Burma et al., 1994); b) nuclear extracts prepared from five different insect cell lines expressing different levels of reporter protein displayed differences in the levels of PPBP binding to the promoter (Mukherjee et al., 1995b); c) the PPBP-promoter complex displayed a different mobility in a gel retardation assay using nuclear extract from a cell line which does not support transcription from the polyhedrin promoter, although the molecular weight of the factor in a UV crosslinking gel was found to be 30-kDa, suggesting the interplay of additional factors along with PPBP in the transcriptionally competent (virus infected) cell lines (Mukherjee et al., 1995b); d) PPBP displayed both duplex promoter DNA binding and also ssDNA binding restricted to the coding strand of the promoter (Mukherjee et al., 1995a) which provided an attractive model to explain repeated rounds of transcription; e) PPBP can also bind to transcriptionally important motifs in another baculovirus very late promoter, the p10 promoter although with critical differences in ssDNA binding and binding to the 5' untranslated regions.

These indirect evidences correlating a host factor with transcription from the polyhedrin promoter is an enigma given the definitive requirement of virus-specific trans-acting factor(s) in this process. In this chapter experimental evidences in support of the involvement of PPBP and PPBP-binding motifs in expression from this promoter in vivo are presented. Further, the involvement of the general transcription factor TATA binding protein (TBP) in polyhedrin transcription is also addressed. All in vivo promoter analyses experiments were carried out using the firefly luciferase reporter which is about 1000-fold more sensitive than the standard CAT assay (deWet et al., 1987; Williams et al., 1989) and as little as $10^{-20}$ moles of luciferase protein can
be detected. The implications of the presence of PPBP-like factors in other eukaryotic systems are discussed.

1. Mutations in the PPBP-Cognate Sequence Motifs Within the Polyhedrin Promoter Abolish Expression from the Polyhedrin Promoter In Vivo

To demonstrate that binding of PPBP to the cognate sequences, essentially constituting the octa- and hexa-motifs within the initiator promoter, is critical for transcription \textit{in vivo}, three sets of 65-mer complementary oligonucleotides with flanking \textit{Hind}III and \textit{Sal}I sites at the 5' and 3' ends, respectively encompassing the polyhedrin promoter (-5 to -65) were designed (Table I). These oligonucleotides either represented the wild type polyhedrin promoter or mutated PPBP-cognate motifs where the hexa-motif (CCGCC in place of AATAAA) or the octa-motif GCCTGCGG in place of TAAGTATT) was altered. The starting 4.6 Kb construct pAJluc used subsequently to clone mutated promoter sequences, was generated by ligating the luciferase reporter gene at the \textit{Bam}HI site of pUC18 (Fig. 4.1, panel A). The luciferase gene was obtained as a 1892 bp \textit{Bam}HI fragment from the baculovirus transfer vector \textit{pAcluc} (Hasnain and Nakhai, 1990). The plasmid construct, pAJluc, was characterized (panel B) using the following restriction enzymes: \textit{Hind}III (lane 3) which linearized the plasmid DNA within the vector, \textit{Eco}RV (lane 4) linearized the construct within the \textit{luc} gene insert, \textit{Bam}HI (lane 5) released the 1.8 Kb \textit{luc} gene, \textit{Sma}I (lane 6) confirmed the orientation of the \textit{luc} gene by linearizing the plasmid DNA. A combination of \textit{Bam}HI and \textit{Hind}III (lane 7) also released the 1.8 Kb \textit{luc} gene since there is no \textit{Hind}III site within the \textit{luc} gene. The same gel was probed with the labeled \textit{luc} cDNA to confirm the presence of \textit{luc} gene (panel C). For the construction of pAJpol\textit{luc} harboring the unmutated polyhedrin promoter, the 92 bp \textit{Eco}RV-\textit{Bam}HI promoter fragment was obtained from the transfer vector pVL1393 (Luckow and Summers, 1989) which was ligated at the \textit{Hinc}II-\textit{Bam}HI site of plasmid vector pUC18 to generate the construct pAJpol. Subsequent to this construction, the luciferase gene was cloned at the \textit{Bam}HI site of pAJpol such that it was driven by the polyhedrin promoter to generate the
Table 1: Oligonucleotides with the promoter mutations. The sequences and the respective plasmid constructs of the 65-mer oligonucleotides with the promoter mutations are described. The mutations have been underlined. The lower case letters depict HindIII overhang which was used for cloning into pAJpol-luc. The 5' end of the complementary oligonucleotide has SalI overhang. pAJpol-luc has the polyhedrin promoter from -1 to -92.
Fig. 4.1 Construction of plasmid construct, pAJluc. (A) Schematic representation of the cloning strategy. All promoter mutations were synthesized as 65-mer complementary oligonucleotides which were annealed and cloned at the HindIII-SalI restriction sites within pAJluc. Lane 1 is uncut plasmid DNA, lane 2 is HindIII digested linearized plasmid, lane 3 is EcoRV digested linearized plasmid DNA, lane 4 is BamHI digestion which releases a 1.8 Kb luc cDNA from the 2.7 Kb pUC18 vector backbone. Lane 5 confirms the orientation of luc within the plasmid construct with SmaI digestion and lane 6 shows the BamHI-HindIII digestion pattern which also releases the 1.8 Kb luc fragment. M is the 1 Kb ladder as a DNA size molecular marker. (C) The same gel was probed with labeled luc cDNA which hybridizes to uncut plasmid DNA (lane 1), linearized DNA (lanes 2, 3, and 5), and only to the released luc fragment without hybridizing to the vector backbone (lanes 4 and 6).
Fig. 4.2 Construction of plasmid construct, pAJpolluc. (A) Schematic to represent the cloning strategy. The 92 bp polh promoter fragment was cloned at the HindII-BamHI sites of pUC18 to generate the construct pAJpol in which the luc cDNA fragment was subsequently cloned at the BamHI site to generate pAJpolluc. (B) Restriction digestion pattern of pAJpolluc. Lane 1 is uncut plasmid DNA, lane 2 is HindIII digested linearized plasmid, lane 3 is EcoRV digested linearized plasmid DNA, lane 4 is BamHI digestion which releases a 1.8 Kb luc cDNA from the 2.8 Kb pAJpol vector backbone. Lane 5 confirms the orientation of luc within the plasmid construct with SmaI digestion. M is the 1 Kb ladder as a DNA size molecular marker. (C) The same gel was probed with labeled luc cDNA which hybridizes to uncut plasmid DNA (lane 1), linearized DNA (lanes 2, 3, and 5), and only to the released luc fragment without hybridizing to the vector backbone (lane 4).
construct pAJpo\textit{lluc} (Fig. 4.2, panel A). The clone was checked with the following restriction enzymes (panel B): \textit{Hind}III (lane 3) linearized the 4.5 Kb construct within the vector, \textit{EcoRV} (lane 4) linearized the construct within the \textit{luc} insert, \textit{BamHI} (lane 5) released the 1.8 Kb \textit{luc} fragment and \textit{SmaI} (lane 6) confirmed the orientation of \textit{luc} with respect to the promoter DNA fragment. The presence of \textit{luc} reporter gene was confirmed by probing the gel with labeled \textit{luc} cDNA fragment in an in-gel hybridization (panel C) reaction.

The complementary 65 mer oligonucleotides carrying the hexa- and octa-mutations were annealed and cloned at the \textit{Hind}III-\textit{Sall} site of pAJ\textit{luc} (Fig. 4.1, panel A) to generate the respective promoter mutant plasmids. The identity of these promoter constructs was confirmed by dideoxy sequencing (Fig. 4.3, panel A). \textit{In vitro} binding analyses of these 65-mer oligonucleotides on a gel mobility shift assay failed to generate PPBP-promoter complex for the mutant derivatives of the promoter (Fig. 4.3, panel B, lanes ) as also reported earlier (Burma et al., 1994).

Following this, the three recombinant promoter fusion constructs were used in a luciferase transient expression \textit{in vivo}. Mutation of the PPBP binding motifs resulted in a near zero (Fig. 4.3, panel C) luciferase expression as compared to the unmutated polyhedrin promoter construct. The inability to detect luciferase expression above the cut-off limit by the very sensitive luminometric assay was true for both the hexa-octa mutations construct (comprising the transcription start point) as well as the hexa mutation construct. Dot blot analysis of the transfected plasmid DNA was also carried out (Fig. 4.3C) to ascertain that equal amounts of the various plasmids had transfected into the insect cells and the observed difference in expression was the result of the function of the involved motifs and not because of the different quantities of the plasmid DNAs. This 'knock-out' data directly demonstrate that in a situation where binding of PPBP is eliminated due to absence of cognate sequence motifs, \textit{in vivo} expression of a reporter gene from the polyhedrin promoter is also abolished.
Fig. 4.3 Hexa-motif and Octa-motif are required for polyhedrin promoter activity. (A) The different constructs carrying the unmutated and mutated polyhedrin promoter are described. The sequencing gels showing the mutations are shown on the left.
Fig. 4.3 (B) Labeled B domain (lanes 1-3), mutHex (lanes 4-6) and mutOct (lanes 7-9) oligonucleotides were incubated either in the absence of protein extract (lanes 1, 4, and 7) or in presence of 1 μg of Sf9 nuclear extract (lanes 2, 3, 5, 6, 8 and 9). The oligonucleotide-protein binding was competed in presence of 25 ng of unlabeled B domain (lane 3), or mutHex (lane 6), or mutOct (lane 9). (C) Transient luciferase expression measured 60 h pi in a luminometer. The values have been normalized with respect to the total amount of plasmid DNA taken up by the insect cells. DNA dot blot analyses demonstrating that equal amounts of plasmid DNA constructs have transfected into the insect cells are shown on the right. 1 and 2 refer to DNA from duplicate wells.
2. Individual Bases Within the Hexa-Motif are Sensitive to PPBP Binding *In Vitro*

An interesting observation from the above experiment was the involvement of the hexa-motif alone in expression from the polyhedrin promoter. The octa-motif involves the transcription initiation site and the loss of expression from the polyhedrin promoter after its mutation is therefore understandable. Having demonstrated the importance of the hexa-motif alone in *in vivo* expression, mutational analysis of this sequence was investigated. Alignment of the polyhedrin and p10 promoter revealed that the hexa-motif shares four bases TAAA between the two promoters at positions -52 to -55 and -72 to -75 with respect to the polyhedrin and p10 promoters, respectively (Chapter III). Since both promoters can bind PPBP and that PPBP-p10 duplex promoter interaction also involves the sequences spanning the hexa-motif, it is conceivable that these last four bases may alone be involved. Four sets of complementary oligonucleotides spanning the polyhedrin promoter from -5 to -65, with restriction sites for *HindIII* and *SalI* at the 5' and 3' ends, respectively to facilitate cloning, were synthesized (Table I). Within these oligonucleotides the hexa-motif AATAAA was mutated to either AAGAAA or AATCAA or AATACA or AATAAAC to give rise to mH3T, mH4A, mH5A and mH6A constructs, respectively. *In vitro* binding analysis of these 65-mer oligonucleotides is shown in Fig. 4.4, panel A. The PPBP-B domain complex (lane 2) and its specificity in a homologous cold competition assay (lane 3) was used as controls. It is apparent that none of the mutations gave rise to a complex similar in mobility to the PPBP complex and instead gave a complex differing in intensity and/or mobility (lanes 5, 8, 11 and 14) although same amount of nuclear extract and equal amounts of the labeled oligonucleotides were used. The complexes obtained with mH3T and mH5A (lanes 5 and 11) were much stronger than those obtained with mH4A and mH6A (lanes 8 and 14). Furthermore, mH3T, mH4A and mH5A complexes were specific as can be seen with 25-fold excess of unlabeled oligonucleotides in cold competition assays (lanes 6, 9 and 12) whereas the mH6A complex seemed to be non-specific since it could not be competed with 25-fold excess of unlabeled mH6A oligonucleotide (lane 15). These
Fig. 4.4 Mutations within individual bases of the hexa-motif affect polyhedrin promoter activity. (A) Labeled 32-mer polhB domain or 65-mer mH3T domain or mH4A domain or ml5A domain or mH6A domain were either incubated alone (lanes 1, 4, 7, 10, and 13) or with 1 μg of Sf9 nuclear extract (lanes 2, 3, 5, 6, 8, 9, 11, 12, 14 and 15). The respective DNA-protein complexes obtained were competed with 25 ng of respective homologous unlabeled domains (lanes 3, 6, 9, and 15).
Fig. 4.4 (B) Plasmid constructs with point mutations within the polyhedrin promoter are shown. (C) Transient luciferase expression from recombinant hexa-motif constructs. *Luc* expression was measured 60 h pi in a luminometer. The values have been normalized with respect to the total amount of plasmid DNA taken up by the insect cells. DNA dot blot analysis, shown on the right, demonstrates that equal amounts of plasmid DNA constructs have transfected into the insect cells. 1 and 2 refer to DNA from duplicate wells.
results demonstrate the sensitivity of the individual nucleotides within the hexanucleotide motif in terms of PPBP-complex formation in vitro.

3. Individual Bases Within the Hexanucleotide Motif are Also Important for Expression of a Luciferase Reporter from the Polyhedrin Promoter In Vivo

The four sets of oligonucleotides carrying point mutations within the hexa-motif were cloned into the promoter minus luciferase expression vector pAJluc in a similar manner as described earlier to drive the expression of the luc reporter gene. The point mutations were confirmed by dideoxy sequencing using the universal M13 forward primer (Fig. 4.4, panel B). These mutated plasmid reporter constructs were used as before for in vivo transient expression assay and luciferase expression levels were measured at 60 h pi (Fig. 4.4, panel C). They were compared with respect to the unmutated construct pAJpolLuc. While appreciable levels of luciferase expression above the cut-off limit was observed for pAJpolLuc there was a drastic reduction in expression with any of the point mutant constructs. These results demonstrate that mutation of any of the bases within the hexanucleotide motif affects PPBP binding and subsequently the in vivo expression of the reporter gene. These observations also reiterate the fact that binding of PPBP to its unmutated cognate motifs is an essential event prior to the polyhedrin promoter driven expression thus demonstrating the importance of PPBP:cognate sequence interaction in transcription from the polyhedrin promoter.

4. PPBP and TATA Binding Protein (TBP) are Different with Respect to their Binding Activities to PPBP-Cognate Motifs

Experiments were designed to determine whether transcription from the polyhedrin promoter also directly involved the ubiquitous factor, the TATA-binding protein (TBP), so important for eukaryotic transcription initiation (Roeder, 1991). Additionally, given the apparent similarities between TBP and PPBP (Burma et al., 1994), has the latter taken over the role of TBP in so far as polyhedrin transcription is
Fig. 4.5 PPBP is not the same as TBP. (A) Labeled TFIID consensus oligonucleotide (lanes 1-8) and polhB domain (lanes 9-12) were used in gel mobility shift assays either alone (lanes 1 and 9), or with 1 μg of HeLa cell nuclear extract (lanes 2 and 10), or with 20 ng of purified TBP (Promega, USA) (lanes 4, 7, 8 and 12), or with 1 μg of Sf9 nuclear extract (lanes 3 and 11). For supershift assay TFIID oligonucleotide was either incubated alone with 1:500 dilution of normal rabbit serum (NRS, lane 5) or 1:500 dilution of anti-TBP serum (lane 6) or with normal rabbit serum and 20 ng of purified TBP (lane 7) or with anti-TBP serum and 20 ng of purified TBP (lane 8). (B) Gel retardation assays were carried out either with labeled polhB domain (lanes 1-4 and 9) or TFIID oligonucleotide (lanes 5-8 and 10). The labeled polhB domain was either incubated alone (lane 1), or with 1 μg of Sf9 nuclear extract (lanes 2-4), or with 1 μg of HeLa cell nuclear extract (lane 9). The DNA-protein complex obtained (lane 2) was competed with 25 ng of unlabeled polhB domain (lane 3), or with 25 ng of unlabeled TFIID domain (lane 4). The labeled TFIID oligonucleotide was either incubated alone (lane 5), or with 1 μg of Sf9 nuclear extract (lanes 6-8), or with 1 μg of HeLa cell nuclear extract (lane 10). The TFIID-Sf9 complex obtained (lane 6) was competed with 25 ng of unlabeled polhB domain (lane 7), or with 25 ng of unlabeled TFIID domain (lane 8).
Fig. 4.5 (C) Gel retardation assays were carried out either with labeled polhB domain in presence of 2 µg (lane 1) or 4 µg of purified PPBP (lane 2) or with TFIID oligonucleotide in presence of 4 µg of purified PPBP (lane 3). (D) Western blot analysis. Lane 1 contains 60 ng of purified SP1 as negative control, lane 2 contains 100 µg of Sf9 nuclear extract, lanes 3 and 4 have 60 ng of purified TBP and 100 µg of HeLa cell nuclear extract respectively as positive controls. The blot has been probed with 1:3000 dilution of anti-hTBP antibody. Protein molecular size markers (in kDa) are shown on the right.
In vivo Role of PPBP-Binding Motifs

concerned? To address these questions EMSAs with a consensus TFIID duplex oligonucleotide (5'GCAGAGCATATAAGGTGAGGTAGGA3') and purified TBP (Promega, USA) were carried out using gel conditions specific for TFIID binding (see Materials and Methods). As expected, HeLa cell extract (1 μg) and purified TBP (20 ng) could bind to the synthetic 25 mer oligonucleotide containing the TFIID cognate sequence (Fig. 4.5, panel A, lanes 2 and 4) but not to the polyhedrin B domain oligonucleotide containing the PPBP-cognate sequence motifs (lanes 10 and 12). Conversely, Sf9 nuclear extract (1 μg) could bind to the B domain to form a PPBP-specific complex (lane 11) and also to the TFIID oligonucleotide to form a different complex (lane 3). The Sf9 extract-TFIID oligo complex was similar in mobility to the HeLa cell extract-TFIID oligo and the purified TBP-TFIID oligo complex, but was distinct from the PPBP-B domain complex. The complexes described above were specific in their binding to the respective domains as seen in panel B. The PPBP-B domain complex (Fig. 4.5, panel B, lane 2) could be specifically competed with a 25-fold excess of unlabeled B domain (lane 3) but not with a similar excess of the unlabeled TFIID oligonucleotide (lane 4). Similarly, the Sf9 extract-TFIID complex (lane 6) could be specifically competed with a 25-fold excess of unlabeled TFIID oligonucleotide (lane 8) but not with an equal amount of unlabeled polyhedrin promoter B domain (lane 7). Under the present binding conditions, HeLa cell extract could not bind the polyhedrin B domain (lane 9) whereas it could do so to the consensus TFIID oligonucleotide to give a DNA-protein complex similar in mobility to the Sf9 extract-TFIID complex (compare lanes 6 and 10). Further, affinity purified PPBP could specifically bind to polyhedrin promoter "initiator" region (Fig. 4.5, panel C, lanes 1 and 2) but could not do so to the TFIID consensus sequence (lane 3). These results clearly demonstrate the presence of a specific TBP-like activity in Sf9 nuclear extract (Rasmussen and Rohrmann, 1994) distinct from PPBP. Western blot analysis (panel C) using a rabbit antiserum against the C-terminal domain of hTBP (a kind gift from Dr. Robert G. Roeder, Rockefeller University, USA) further confirmed the presence of a TBP like activity in Sf9 cell nuclear extract. Anti-TBP antiserum at a dilution of 1:3000, as expected, gave a specific 37.7 kDa band when crude HeLa cell nuclear extract and purified TBP
(Promega, USA) were used as positive controls (lanes 4 and 3). Negative control using purified Sp1 (Promega, USA) failed to generate any detectable signal (lane 1). A specific ~30 kDa band was seen with the Sf9 nuclear extract (lane 2). A similar blot probed with pre-immune rabbit serum did not give any signal (data not shown). As expected, any supershifting of the consensus TFIID sequence-purified TBP complex with 1:500 dilution of anti-TBP antibody (panel A, lane 8) could not be detected since these antibodies are against the C-terminal domain of TBP which is involved in binding to the DNA. These results while demonstrating the presence of distinct TBP-like activity in Sf9 cells also strongly argue against TBP having a direct role in polyhedrin promoter driven transcription since it is unable to directly contact the polyhedrin "initiator". These observations also unequivocally distinguish TBP from PPBP, the initiator binding protein.

5. PPBP-Like Activity is Present in Other Eukaryotic Systems

Experiments were designed to determine if PPBP or PPBP-like factors are present in other eukaryotic systems. Nuclear extracts from HeLa cells, pea nuclei and chloroplasts were prepared and used in electrophoretic mobility shift assays using labeled B domain of the polyhedrin promoter which contains the specific PPBP-binding motifs viz. the hexa-motif (AATAAA) and the octa-motif (TAAGTATT). EMSAs carried out (using the Sp1-specific gel conditions as described in Materials and Methods) with the B domain of the polyhedrin promoter and Sf9 nuclear extract generated the PPBP-specific complexes (Fig. 4.6, panel A, lane 2). These complexes were specifically competed out with a 25-fold excess of unlabeled oligonucleotides possessing PPBP-binding motifs (lane 3) but not when oligonucleotide lacking the PPBP-binding motifs (Sp1 containing oligonucleotide or AcSp1 oligonucleotide) were used (lanes 4 and 5). Three complexes were obtained with the HeLa cell extract and the labeled polyhedrin promoter B domain (lane 6) which could again be specifically competed out with a 25-fold excess of unlabeled B domain (lane 7) but not with a 25-fold excess of non-specific oligonucleotides (lane 8 and 9). At least one of these
**Fig. 4.6 PPBP-like factor is present in HeLa cell nuclear extracts.** (A) Gel retardation assays were carried out in presence of labeled polhB domain in presence of 2 μg of Sf9 nuclear extract (lanes 2-5) or HeLa cell nuclear extract (lanes 6-9). As a negative control, the polhB domain was incubated alone in the absence of any nuclear extract (lane 1). The DNA-protein complexes obtained with the Sf9 nuclear extract (lane 2) and the HeLa cell nuclear extract (lane 6) were specifically competed out with 25 ng of unlabeled polhB domain (lanes 3 and 7), or Sp1 consensus oligonucleotide (lanes 4 and 8), or AcSp1 domain (lanes 5 and 9).
**In Vivo Role of PPBP-Binding Motifs**

Fig. 4.6 PPBP-like factor is present in pea nuclear extracts. (B) Labeled polhB domain was either incubated alone (lane 1), or 1 μg of Sf9 nuclear extract (lane 2), or 1 μg of pea nuclear extract (lanes 3-5), or 1 μg of pea chloroplast extract (lanes 6-9), or 1μg each of Sf9 nuclear extract and pea chloroplast extract (lanes 10-13). Bi and Bii are the same gel where Bii represents a much higher exposure time. The DNA-protein complex obtained with the pea nuclear extract (lane 3) was competed with 25 ng of unlabeled polhB domain (lane 4), or TFIID oligonucleotide (lane 5). Similarly the complexes obtained with the pea chloroplast extract (lane 6) and with a mixture of Sf9 nuclear and pea chloroplast extracts (lane 10) were competed in presence of 25 ng of unlabeled B domain (lanes 7 and 11), or TFIID oligonucleotide (lanes 8 and 12), or non-specific DNA, pUC18 (lanes 9 and 13).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>B Domain (1ng)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sf9 Nuclear</td>
<td>- + - - - - - - + + + +</td>
</tr>
<tr>
<td>Pea Nuclear</td>
<td>- + + + - - - - + + + +</td>
</tr>
<tr>
<td>Pea Chloroplast</td>
<td>- - + + + + + + + + + +</td>
</tr>
<tr>
<td>Competitor</td>
<td>- - B IID - B IID pUC - B IID pUC</td>
</tr>
</tbody>
</table>

**Table 4.6**

<table>
<thead>
<tr>
<th>Substrate</th>
<th>B Domain (1ng)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sf9 Nuclear</td>
<td>- + - - - - - - + + + +</td>
</tr>
<tr>
<td>Pea Nuclear</td>
<td>- + + + - - - - + + + +</td>
</tr>
<tr>
<td>Pea Chloroplast</td>
<td>- - + + + + + + + + + +</td>
</tr>
<tr>
<td>Competitor</td>
<td>- - B IID - B IID pUC - B IID pUC</td>
</tr>
</tbody>
</table>

**Table 4.6**
complexes matched in mobility with the PPBP complexes obtained with the Sf9 nuclear extract. Similar EMSAs (panel B) with the pea nuclear extract and the pea chloroplast extract using labeled B domain oligonucleotide were carried out. A single PPBP specific complex was obtained when the Sf9 nuclear extract was used as a control (lane 2). A specific complex could also be obtained with the pea nuclear extract (lane 3) and the pea chloroplast extract (lane 6) which could be specifically competed out by 25-fold excess of unlabeled B domain (lanes 4 and 7) but not by 25-fold excess of unlabeled TFIID domain (lanes 5 and 8) or by non-specific DNA such as pUC18 (lane 9). Interestingly, these complexes were of a different mobility than those obtained with the Sf9 nuclear extract. Upon mixing the pea chloroplast extract and the Sf9 extract two complexes (lane 10) both of which could be specifically competed out with the unlabeled B domain (lane 11) but not with unlabeled TFIID (lane 12) or unlabeled pUC18 (lane 13). These results clearly demonstrated that a PPBP-like factor was ubiquitously present in diverse eukaryotic systems investigated viz. mammalian HeLa cell extract and the pea extracts. Furthermore, this PPBP-like factor was distinct from the ubiquitous transcription factor (TBP). The PPBP-like factor showed cognate sequence recognition properties characteristic for PPBP.

DISCUSSION

PPBP IS AN INITIATOR BINDING PROTEIN

Many cellular genes do not contain a canonical TATA box-like sequence and are far more simple in their organization. Smale and Baltimore (1989) identified a 17 bp sequence motif encompassing the transcription start site of the murine lymphocyte specific terminal deoxynucleotidyltransferase (Tdt) gene which was termed as the "initiator" and was sufficient for accurate basal transcription both \textit{in vitro} and \textit{in vivo}. Several initiator containing promoters, which probably represent the simplest promoters known so far, have now been identified (Chen et al., 1994; Yoo et al., 1991). A core tetra nucleotide motif CAGT has been located at the RNA start site of the transregulator gene \textit{ie-1} of AcNPV (Pullen and Friesen, 1995) which is involved in its transcription and also resembles the proposed consensus for arthropod
transcriptional initiator elements (Cherbas and Cherbas, 1993). Functional analyses of 80 random and mutant initiator elements (Javahery et al., 1994) identified a loose consensus sequence (Py Py A⁺¹ N T/A Py Py) for such initiator promoter elements. The 18 bp sequence surrounding the baculovirus polyhedrin gene transcription initiation point, comprising of the hexa (AATAAA) and octanucleotide (TAAGTATT) motifs, represents an initiator-like sequence and also shares considerable homology with the initiator consensus sequence (Fig. 1.1).

Several initiator-binding proteins (IBPs) have been identified (Weis and Reinberg, 1992) but the exact mechanism of transcription from such promoters is yet to be worked out. The most well understood IBPs include TFII-I (recognizing the Ad2 IVa2 and Ad2 ML initiators; Roy et al., 1993), YY1 protein (which recognizes the Ad2-associated virus P5 initiator; Weis and Reinberg, 1992) and HIP1 (recognizing the dihydrofolate reductase initiator; Means et al., 1992). Other than directly binding the initiator element and thereby affecting transcription, the IBPs have little in common structurally. The initiator binding protein, YY1, is an extensively characterized transcription factor which binds to the regulatory regions of many viral and cellular genes where it has been reported to function either as a transcriptional repressor (Hariharan et al., 1991) or a transactivator (Seto et al., 1991). A core consensus (5’GCCATNTT3’) for the binding site of YY1 has been reported. This factor is known by various names viz. NF-E1, UCRBP, δ and consists of two proteins of molecular mass 68-kDa and 40-kDa (Becker et al., 1994). YY1 acts as a phosphoprotein and the cellular binding activity is abolished by phosphatase treatment. This protein contains a highly charged domain, part of which is composed of stretches of histidine residues which are found in a number of transcription factors and may be important for function. PPBP also contacts the DNA through histidine residues (Hasnain et al., manuscript in preparation). The reported target sequences specific for this protein vary in nucleotide sequences, although the majority of them appear to be A+T rich. It is pertinent to iterate that the baculovirus initiator element is also A+T rich.

The host factor, PPBP, from Sf9 insect cells which displayed unusual characteristics with respect to affinity, specificity, stability and cognate sequence
requirements is an initiator binding protein which recognizes A-T rich sequences within the AcNPV polyhedrin and p10 promoters. Since PPBP has several characteristics similar to YY1 including the requirement for phosphorylation for target binding, the YY1 consensus was compared to the polyhedrin and p10 promoter sequences to determine if PPBP recognized the same sequence motif as YY1. The absence of any significant homology specifically within the initiator regions of p10 and p29 promoters and the YY1 consensus makes it unlikely that PPBP and YY1 are similar. Further, PPBP possesses such unusual characteristics (salt tolerance, affinity, specificity and dual ss and ds DNA binding activities; Burma et al., 1994; Mukherjee et al., 1995a) which have not been reported for any known transcription factor. It is conceivable that PPBP as an initiator-binding protein binds the initiator complex which serves to recruit the subsequent players culminating in the formation of transcription initiation complex. PPBP therefore, represents a rare example of an IBP with distinct and characteristic properties.

SIMILARITY AND DIFFERENCES OF PPBP WITH HMG PROTEINS

It is interesting to highlight the similarities and differences between PPBP and the high mobility group (HMG) proteins. The high mobility group domain is a DNA binding motif which is shared by abundant non-histone components of chromatin (Grosschedl et al., 1994). Common properties of HMG domain proteins include interaction with the minor groove of the DNA helix, binding to irregular DNA structures and the capacity to modulate DNA structure by bending. HMG-like proteins preferentially bind to AT-rich promoter elements (Reeves and Nissen, 1990; Struhl, 1985). It is pertinent to compare the apparent similarities between PPBP and HMG proteins. PPBP also binds to a specific AT-rich sequence and not to any stretch of AT-rich sequences. It also binds DNA through the minor groove. It is also conceivable that PPBP has a HMG-like domain to contact DNA and a transcription activation domain to contact other factor(s) for transcription initiation through protein-protein interactions. Quite recently, certain insect proteins have been shown to be structurally homologous to mammalian high mobility group proteins I/Y (Claus et al.,
In Vivo Role of PPBP-Binding Motifs

1994). Several other features of PPBP, however make it distinct from classical HMG proteins. Although it has ssDNA binding activity like few other HMG proteins (Isackson et al., 1979) none of the latter has been shown to switch in between the two binding activities the way PPBP does. It is essential for PPBP to be phosphorylated to bind its cognate motifs (Burma et al., 1994) whereas it is quite the opposite for the known HMG proteins (Reeves et al., 1991; Nissen et al., 1991). Apart from this, no known HMG protein has properties characteristic for PPBP. This then makes PPBP a transcription factor with a difference which shares features with some known proteins as well as unrelated ones while still maintaining its own identity.

POSSIBLE MODE OF ACTION OF PPBP

Subtractive hybridization or marker rescue experiments have identified several lef, vlf, and hcf-1 gene products encoded by the viral genome (McLachlin and Miller, 1994; Todd et al., 1995; Lu and Miller, 1995a; Lu and Miller, 1995b). These genes may be directly or indirectly involved in polyhedrin promoter activation and must function in conjunction with cellular factors. Direct interaction and involvement in transcription from late and very late promoters for none of these factors has been demonstrated so far. PPBP therefore, constitutes the only host factor that binds to transcriptionally important motifs within the baculovirus polyhedrin promoter and is a likely candidate to be involved in cross-talks with other accessory transcription factors including the lefs. In the simplistic model implicating PPBP in polyhedrin transcription reconciling PPBP’s duplex promoter binding activity with a half life of 15 min and sequence specific ssDNA binding activity with a half life of 60 min, this host factor after scanning the promoter binds to the initiator element. Having recruited other factors, perhaps via the TBP and/or including virus-specific factors (lefs?), it then switches over to a ssDNA binding regime, binding specifically only to the coding strand of the promoter keeping it in place (longer half life) for repeated rounds of transcription. The observation of a putative helicase activity and DNA dependent ATPase activity in PPBP (Hasnain et al., manuscript in preparation), a characteristic shared between PPBP and TFIIH (Schaeffer et al., 1993;
In Vivo Role of PPBP-Binding Motifs

Serizawa et al., 1993; Drapkin and Reinberg, 1994), renders possible the unwinding of the DNA required between the ds and ss DNA binding events without PPBP moving away from the DNA. These results fit well with a proposed model (Martinez et al., 1994) for transcription initiation from initiator containing promoters through an initiator binding protein thereby rendering PPBP an important component of the transcription initiation complex. According to this model TBP is tethered to a TATA-containing promoter by direct binding to the TATA box and promotes the assembly of a functional pre initiation complex. In contrast, the TATA-binding activity of TBP is not required and direct TBP-DNA interactions through domains distinct from the TATA-binding domain are not sufficient for the functional recruitment of TBP to a 'true' TATA-less promoter. For initiator-dependent 'true' TATA-less promoters, one or several TAFs are required for the functional recruitment of TBP to the core promoter region, either by direct interactions with the initiator and its downstream region or by interaction with an initiator binding protein (IBP). That TBP is present in Sf9 nuclear extracts has already been shown (this Chapter) but it is unable to directly contact the polyhedrin initiator. In such a situation and given the fact that PPBP is an initiator binding protein which directly sits at the initiator, it is conceivable that this promoter also utilizes a similar mechanism for the formation of transcription initiation complex. Another point to be considered is that interaction of PPBP with the double helix, like TBP (Lee et al., 1991) and other transcription factors such as human cytomegalovirus IE2 protein (Lang and Stamminger, 1994) and HMG-IY protein (Solomon et al., 1986), is through the minor groove (Hasnain et al., manuscript in preparation). The important question however, remains as to what other factor(s) PPBP interacts with.

The above scenario fits with the "sequential assembly" model (Buratowski, 1994) for initiation complex formation. Unlike the newly emerging alternative "holoenzyme" model, in the "sequential assembly" model the general transcription factors and polII assemble into an initiation complex in a stepwise fashion. In the other model (Koleske and Young, 1994; Koleske and Young, 1995) the assembled holoenzyme (polII with other general transcription factors) is recruited to promoters
at which TFIID is already bound and serves as the "landing pad" for the holoenzyme. The holoenzyme is highly stable even in the absence of DNA. Even in this model, the template is recognized by the TFIID complex consisting of TBP and the associated TAFs and in cases where TBP does not directly contact the promoter DNA, by the IBP. Whatever be the 'order' of transcription initiation complex formation, the important role of PPBP in initiation complex formation at the polyhedrin initiator can not be undermined.

ON THE NATURAL FUNCTION OF PPBP WITHIN THE HOST CELL

The fact that the virus recruits a host factor for transcribing one of its most important viral genes then raises the other fundamental question about the function of PPBP *per se* within the insect cell. The possibility of the ATAAG motif of the polyhedrin promoter originating from the host genome (Friesen et al., 1986) readily explains why this motif is recognized by a host factor. PPBP also recognizes the motif AATAAAA which very interestingly is identical to the polyadenylation recognition sequence (Wahle and Keller, 1992). This hexa-motif in the polyhedrin and the p10 initiator promoter is however, not a polyadenylation signal in the context of the polyhedrin or the flanking open reading frames (Ayres et al., 1994) on the viral genome. Putative hexa-motifs corresponding to AATAAAA are also spread throughout the AcNPV sequence (Westwood et al., 1993) but do not act as transcription start sites unless they are in the neighborhood of the transcription start point as in the case of the polyhedrin promoter. Fractionation of the HeLa cell extracts for dissecting the polyadenylation machinery described factors required for cleavage and/or polyadenylation. These include two cleavage factors, a cleavage-polyadenylation specificity factor (CPSF), a cleavage stimulation factor (CstF) and poly(A) polymerase. Purified CPSF consists of four polypeptides one of which is ~30 kDa in molecular weight which previously escaped detection (Jenny et al., 1994). Is it possible that PPBP belongs to the CPSF complex of the polyadenylation machinery? Several lines of evidences indicate such a possibility. That PPBP is also a 30 kDa protein with sequence specific ssDNA binding activity and AATAAAA as the recognition motif-
characteristics it shares with the 30-kDa CPSF subunit- appears to be more than coincidental. The sensitivity of individual nucleotides within the hexa-motif to PPBP binding and in vivo function demonstrated earlier in this chapter is yet another feature shared between PPBP and the 30-kDa CPSF factor (Bardwell et al., 1991; Wickens, 1990). EMSAs using PPBP cognate motif oligonucleotide and nuclear extracts prepared from HeLa cells and pea chloroplast and nuclei indeed revealed the presence of PPBP-like factor under PPBP-specific gel conditions- an observation which can be reconciled only if PPBP has a more fundamental role across the eukaryotes. The binding of PPBP present in the insect cell nuclear extract and that in the pea nuclear extract to the B domain is mutually exclusive as seen from mixing experiments where the two extracts since two distinct complexes are obtained. Had this not been the case, a distinct third complex with a much slower mobility would have been obtained.

To summarize, this chapter details experimental observations which unequivocally demonstrate that PPBP-binding motifs and the consequent binding of PPBP constitute an essential event for detectable expression from the polyhedrin promoter in vivo. Further, PPBP is an initiator binding protein very distinct from the TATA binding protein. It is conceivable that this host factor present in a number of other systems is involved, normally, in a more fundamental process of polyadenylation. The virus through, yet unknown mechanism(s), recruits this host factor to perform transcription of the very important polyhedrin gene.